CRISPR-Cas9 and ideas for Cotton (*Gossypium spp.*)

Dr. Ishwarappa, S. Katageri

Associate Director of Research

University of agricultural Sciences

Regional Research Station, Vijayapura

Karnataka State

katageriis@uasd.in

Genome Editing: Why

- Functional Validations of gene
- Targeted trait modification

Genetic manipulation in crops: History

- Crop Domestication : Natural Genetic Variability (Spontaneous Mutations)
- Crop Domestication : 10,000 years back, about 2500 crop species domesticated
- Conventional Breeding
- Genetic Engineering: GM crops, Molecular Breeding
- **Targeted Gene editing**

Induced Mutations

- Earliest Genome Editing example in plants by Stadler, 1928, **Barley,X-Rays**
- Mari (1952)….X-ray… Early flowering Mari barley released in 1962
- 1969: Special training course on mutation breeding (FAO/ IAEA): Breeders tool box

(www.mvd.iaea.org)-mutatnt variety data base

Targeted Gene Editing- Challenge

- Challenge was the design of chimeric nucleases
- To catalyzes the cleavage of DNA, and the second is capable of selectively binding to specific nucleotide sequences of target molecule, providing the nuclease action to this site

Gene editing technologies

- Transcription activator-like effector nucleases (TALENs)
- Zinc-finger nucleases (ZFNs)

- CRISPR-Cas systems-**C**lustered **R**egularly **I**nterspaced **S**hort **P**alindromic **R**epeats (**CRISPR**)-**C**RISPR **as**sociated (**Cas**) : ***creating a buzz*** *in the science world* –Transgene Free

The CRISPR-Cas9 system **currently stands out** as the **fastest**, **cheapest** and most **reliable** system for 'editing' genes.

CRISPR Mechanism

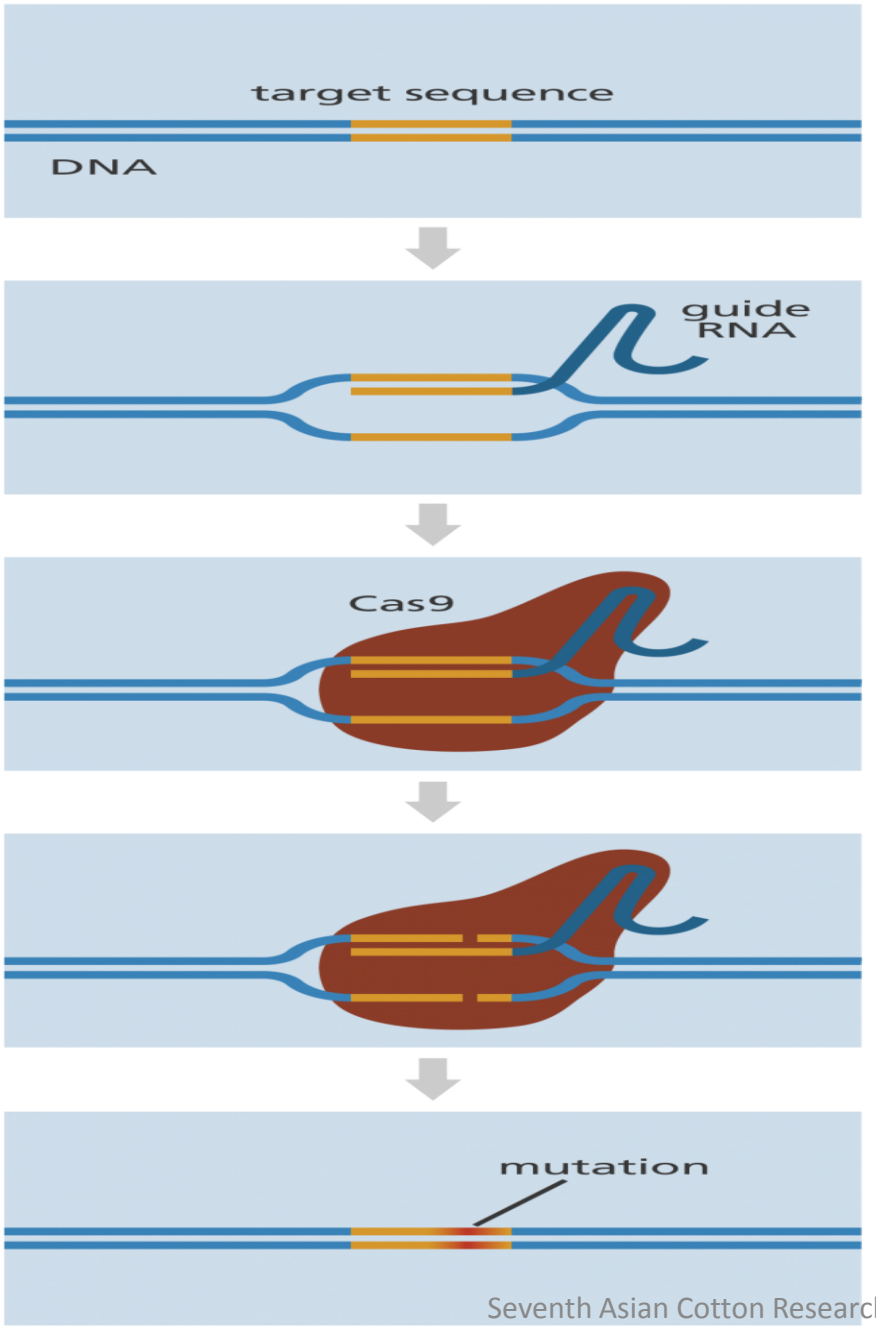
- The key step in editing an organism's genome is selective targeting of a specific sequence of DNA.
- Two biological macromolecules, the Cas9 protein and guide RNA, interact to form a complex that can identify target sequences with high selectivity.

From Where it is learnt

- CRISPR/Cas system is a prokaryotic immune system (form of acquired immunity)that confers resistance to foreign genetic elements
- **CRISPR-**DNA sequences in bacteria that contains snippets of DNA from viruses that have attacked the bacterium
- Snippets are used by the bacterium to detect and destroy DNA from further attacks by similar viruses
- From this basis of a genome editing technology ,known as **CRISPR/Cas9** that allows permanent modification of genes within organisms

How does it work?

- The CRISPR-Cas9 system consists of **two key molecules** that introduce a change (mutation?) into the DNA.
- **Enzyme , Cas9-** acts as a pair of 'molecular scissors' that can cut the two strands of DNA at a specific location in the genome so that bits of DNA can then be added or removed.
- **A piece of RNA -** guide RNA (gRNA)- consists of a small piece of **pre-designed RNA sequence (about 20 bases long)** located within a longer RNA scaffold.



Guide RNA binds to target sequence

Cas9 enzyme binds to guide RNA

Cas9 enzyme cuts both strands of DNA

The cut is repaired introducing mutation

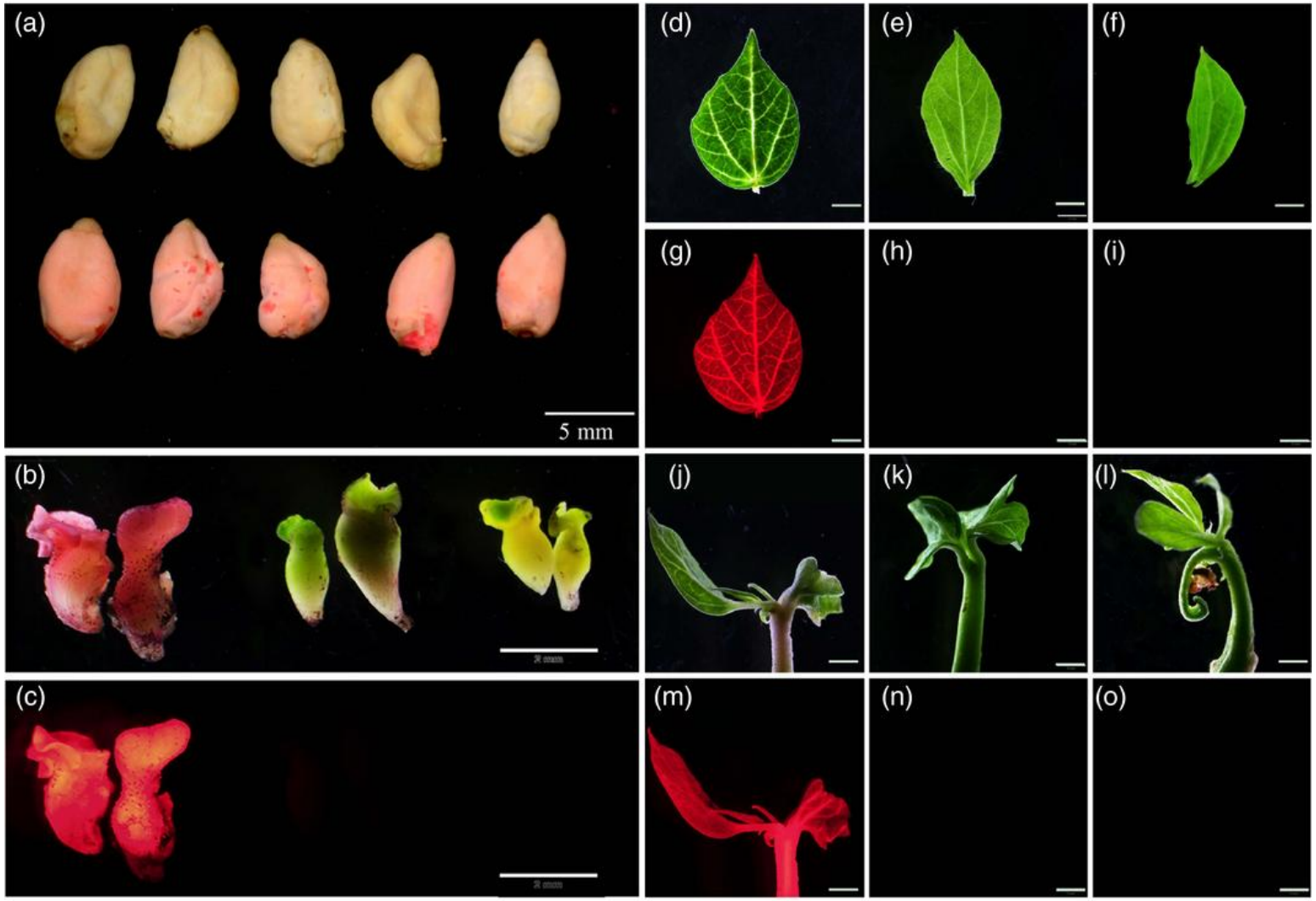
Successful Evidences

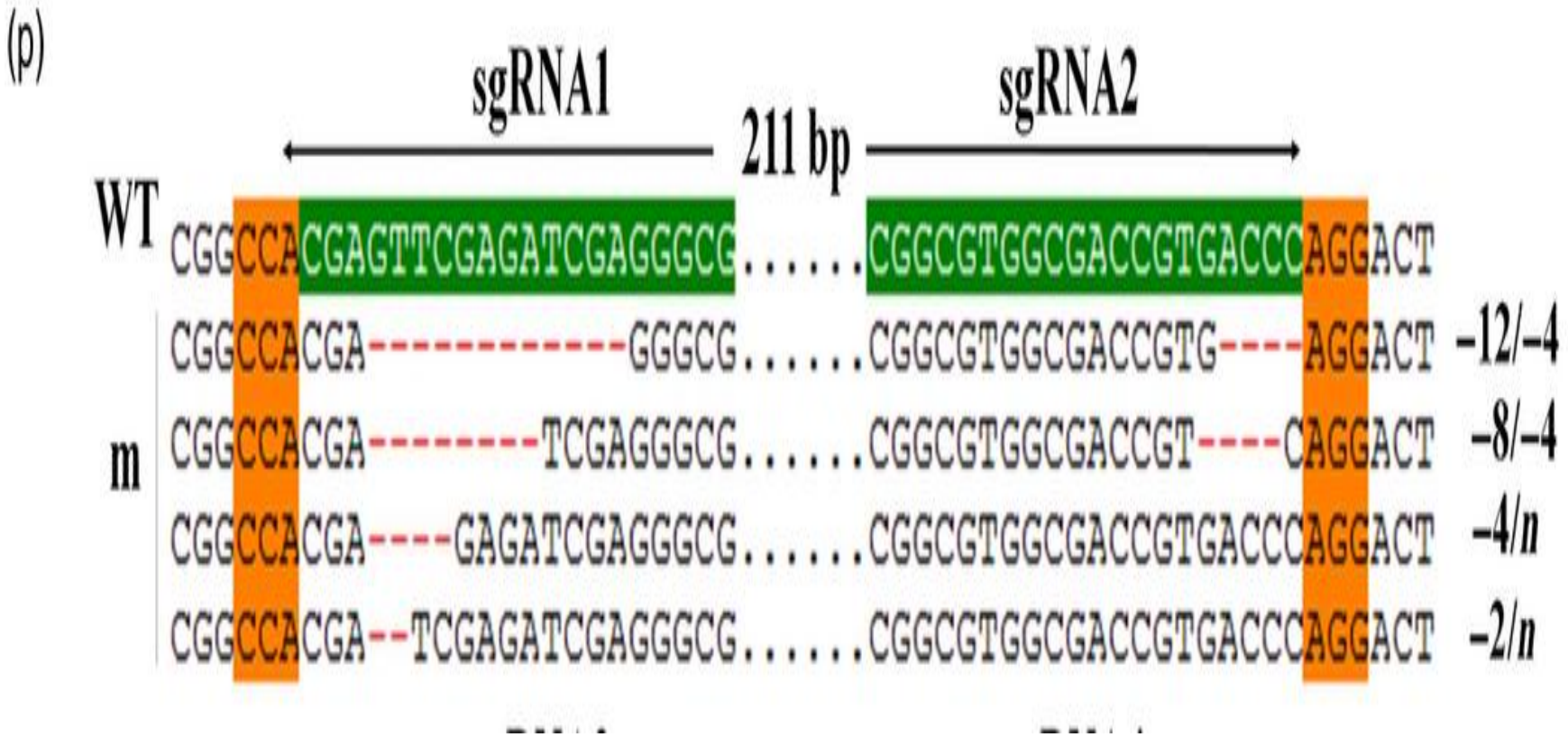
- Cong *et al*. (2013) successfully accomplished the first genome editing in **human cells** with CRISPR/Cas9 system.
- , Mali *et al*. (2013) conducted gene editing in different **human cell** types and accomplished DNA replacement with a donor template.
- **Rice and wheat were the first crops** that were genetically edited with CRISPR/Cas9 system, and knockout of *OsPDS* generated albino rice mutant (Shan *et al*., 2013). These pioneering works promote its broad utilization of CRISPR/Cas9 system in the research of life science.

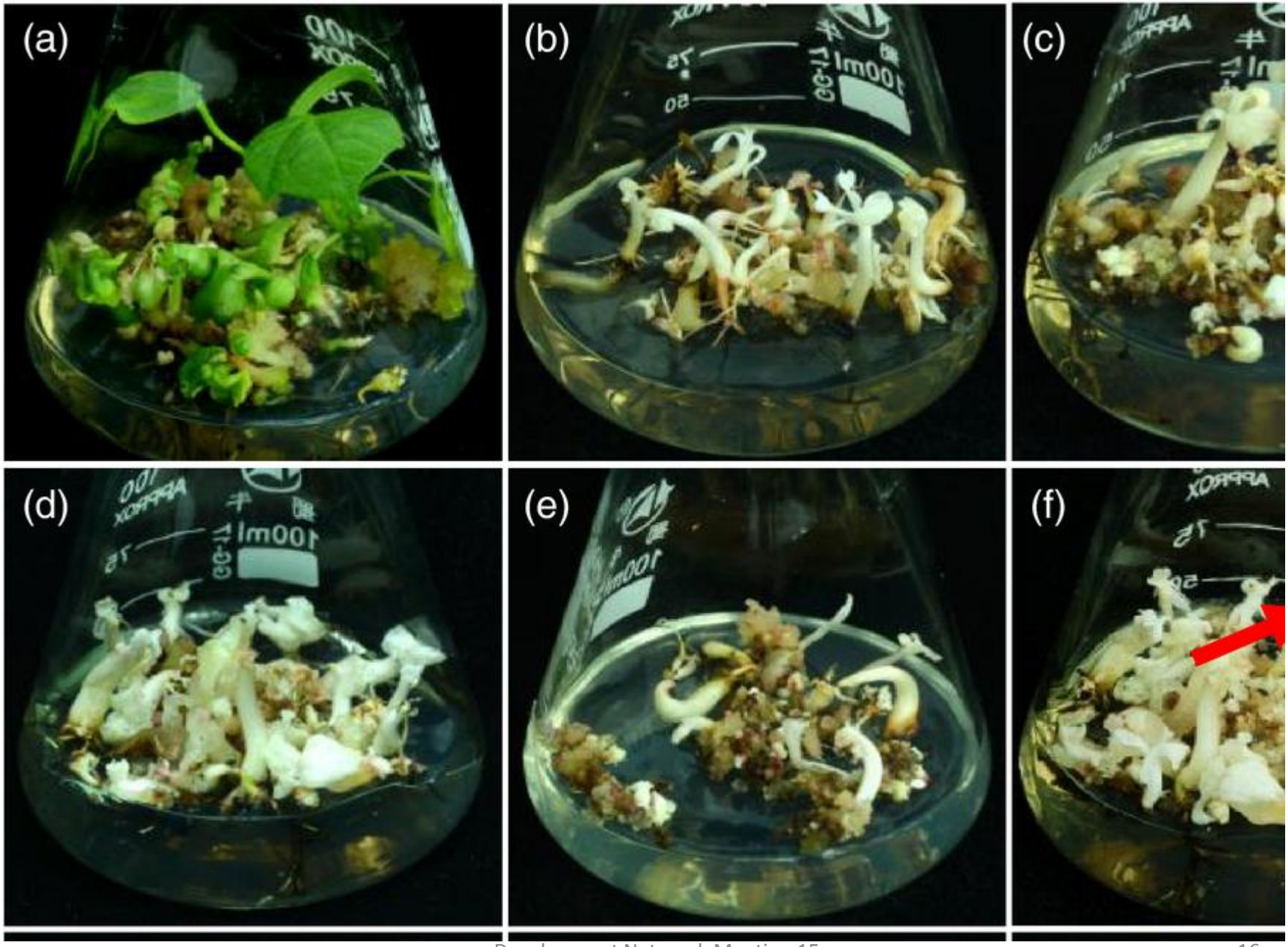
Opportunities in Cotton

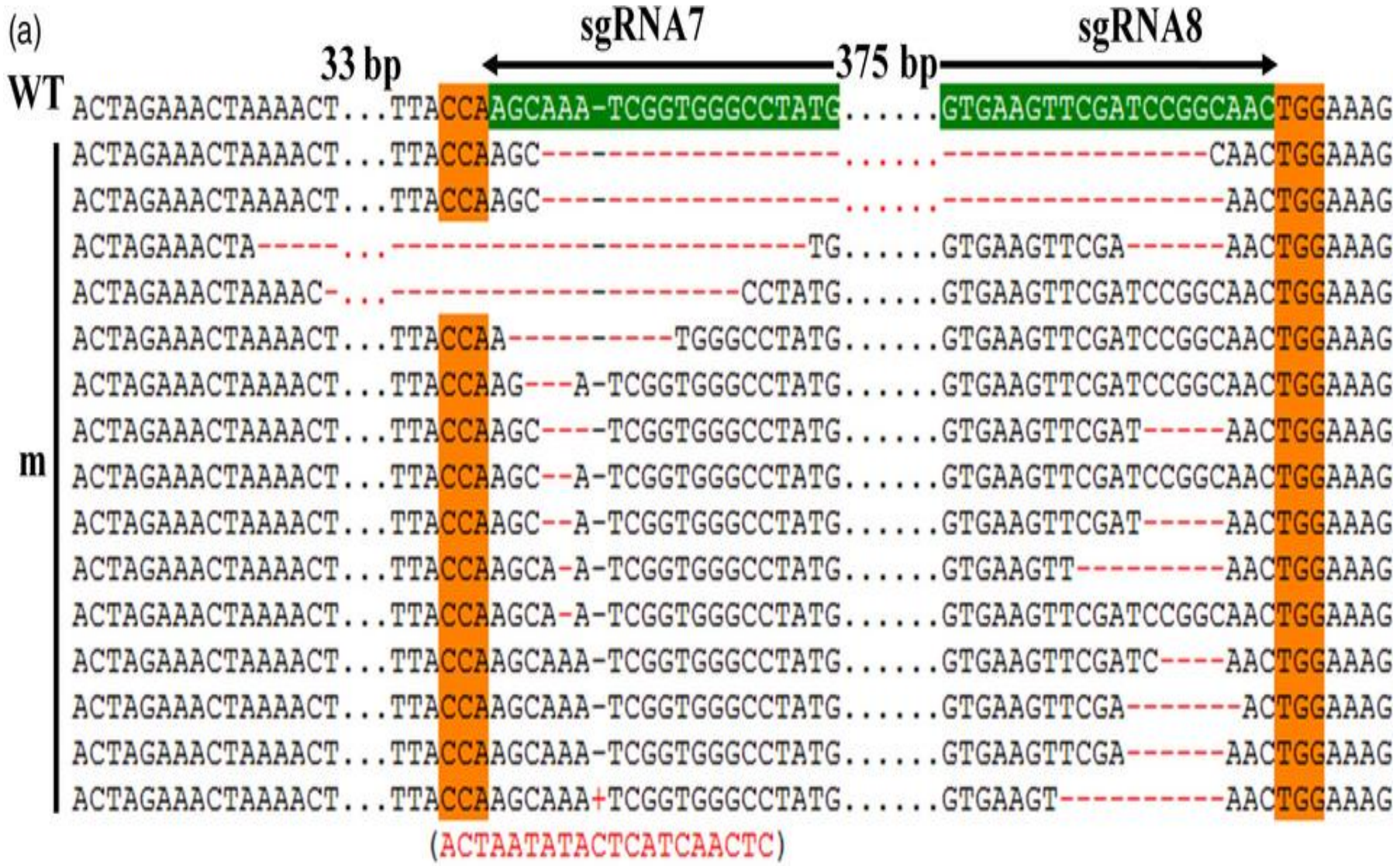
- **High efficient multisites genome editing in allotetraploid cotton (*Gossypium hirsutum*) using CRISPR/Cas9 system**
- Pengchang et al 2017
- *DsRed2* protein (*Discosoma* red fluorescent protein2) was firstly isolated from reef corals (*Discosoma* sp.), and it has been applied in plant molecular biology as a reporter

- *DsRed2*-overexpressed transgenic cotton line-red colour in somatic embryos and seeds under white light
- *DsRed2*-edited T0 plants obtained authentic gene mutation, and they had no red fluorescence
- For another target gene *GhCLA1 (*chloroplast development), an average of 75% regenerated T0 plants showed an albino phenotype


(a)
5 mm
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)
(k)
(l)
(m)
(n)
(o)


(p)
sgRNA1
211 bp
sgRNA2
WT
CGGCCACGAGTTCGAGATCGAGGGCG.....CGGCGTGGCGACCGTGACCCAGGACT
m
CGGCCACGA----------GGGCG.....CGGCGTGGCGACCGTG----AGGACT -12/-4
CGGCCACGA--------TCGAGGGCG.....CGGCGTGGCGACCGT----CAGGACT -8/-4
CGGCCACGA----GAGATCGAGGGCG.....CGGCGTGGCGACCGTGACCCAGGACT -4/n
CGGCCACGA--TCGAGATCGAGGGCG.....CGGCGTGGCGACCGTGACCCAGGACT -2/n


(a)
(b)
(c)
(d)
(e)
(f)


(a)
sgRNA7
sgRNA8
33 bp
375 bp
WT
m
ACTAGAAACTAAAACT...TTACCAAGCAAA-TCGGTGGGCCTATG.....GTGAAGTTCGATCCGGCAACTGGAAAG
ACTAGAAACTAAAACT...TTACCAAGC----------.....----------CAACTGGAAAG
ACTAGAAACTAAAACT...TTACCAAGC----------.....----------AACTGGAAAG
ACTAGAAACTA-----...----------TG.....GTGAAGTTCGA------AACTGGAAAG
ACTAGAAACTAAAAC-...----------CCTATG.....GTGAAGTTCGATCCGGCAACTGGAAAG
ACTAGAAACTAAAACT...TTACCAA----------TGGGCCTATG.....GTGAAGTTCGATCCGGCAACTGGAAAG
ACTAGAAACTAAAACT...TTACCAAG---A-TCGGTGGGCCTATG.....GTGAAGTTCGATCCGGCAACTGGAAAG
ACTAGAAACTAAAACT...TTACCAAGC----TCGGTGGGCCTATG.....GTGAAGTTCGAT-----AACTGGAAAG
ACTAGAAACTAAAACT...TTACCAAGC--A-TCGGTGGGCCTATG.....GTGAAGTTCGATCCGGCAACTGGAAAG
ACTAGAAACTAAAACT...TTACCAAGC--A-TCGGTGGGCCTATG.....GTGAAGTTCGAT-----AACTGGAAAG
ACTAGAAACTAAAACT...TTACCAAGCA-A-TCGGTGGGCCTATG.....GTGAAGTT---------AACTGGAAAG
ACTAGAAACTAAAACT...TTACCAAGCA-A-TCGGTGGGCCTATG.....GTGAAGTTCGATCCGGCAACTGGAAAG
ACTAGAAACTAAAACT...TTACCAAGCAAA-TCGGTGGGCCTATG.....GTGAAGTTCGATC----AACTGGAAAG
ACTAGAAACTAAAACT...TTACCAAGCAAA-TCGGTGGGCCTATG.....GTGAAGTTCGA-------ACTGGAAAG
ACTAGAAACTAAAACT...TTACCAAGCAAA-TCGGTGGGCCTATG.....GTGAAGTTCGA------AACTGGAAAG
ACTAGAAACTAAAACT...TTACCAAGCAAA+TCGGTGGGCCTATG.....GTGAAGT----------AACTGGAAAG
(ACTAATATACTCATCAACTC)

Work done so far.. .. in Cotton

SN	Target gene	Gene function/role	Ref.
1	*GhCLA1*	Chloroplastos alterados	Chen et al. 2017; Wang et al. 2017 Gao et al. 2017
2	*GhVP*	Vacuolar H+ pyrophosphatase	Chen et al. 2017
3	*GhMyb25*	Myb25 like gene	Li et al. 2017
4	*GhARG*	Lateral root formation	Wang et al. 2017
5	*GhEF1*	Translation elongation factor 1	Gao et al. 2017
6	*GhPDS*	Phytoene desaturase to visibly show the effect of genome editing	Gao et al. 2017
7	*Helioithis-Caderin*	Receptor of *H. armigera for Cry1Ac toxin protein*	Wang et al. 2016
8	*GFP*	Green florescent protein as an example of exogenous gene editing in cotton	Junga et al., 2017

Potential target genes in Cotton

SN	Target gene	Gene function/role	Comments
1	*Virus genes (exogenous)*	Virus replication, movement	Exogenous gene target
2	*GhPRP5* *	Proline rich protein, a cell wall structure protein – its expression negatively correlated with fiber quality	To validate the role in fiber length
3	GhCDS*	D-Cysteine desulphydrase –its expression negatively correlated with fiber quality	To validate the role in fiber length
4	GhTLP*	Thaumatin Like Protein-its expression is positively correlated with fiber length	To validate the role in fiber length
5	Other QTLs	Various gene association studies identify several QTLs associated with different traits..	Role of QTLs can be assessed by editing them

CRISPR-Cas9 system to develop gossypol free seed in upland cotton (*G. hirustum* L.)

- More than 90% of cotton produced in India and worldwide is from upland cotton (*Gossypium hirustum* L.). In India every year **2.5 million tons of cotton seed** with **33% protein** **and** **39% fat** content worth **Rs. 10,000 crores** is available.

- Due to the presence of gossypol and other pigments it **is not suitable for human consumption**. Gossypol is desirable in leaves and roots of the cotton plant to control insect pests and diseases.

- Natural **cultivars without gossypol are susceptible to pests/diseases** and hence commercially not successful.

- In USA using **RNAi gossypol** in cotton seed was reduced with a seed specific promoter by suppressing (+)-delta-cadinene synthase (**cdn**) gene with **transgene present** always.

- To develop **transgene free** final product we propose to edit **cdn gene(s) expressed in cotton seeds** using CRISPR-Cas9 and eliminate the CRISPR-sgRNA-Cas9 transgenes in T_2 generation

Selection of sgRNAs

Based on *in silico* analysis selected sgRNAs (single guide RNAs) for each of three cdn genes in *G. hirsutum* L.

G. hirsutum **cdn gene**	**sgRNA sequence (5'- -3') with NGG PAM sequence**	**sgRNA orientation**
cdn1-D1 **(AY800107)**	GTTCTTAAGGGGAAATTATTAGG	Sense
	CCGGTCACTCTTTGAAAATTACC	Anti-sense
	AGGAGAGAAGATGATTTCTCGGG	Sense
	CCTGGAAGGATGTGAATAAAGGG	Anti-sense
	GTTCTTAAGGGGAAATTATTAGG	Sense
	CCAACCGAAATGCCAATAGAAGT	Anti-sense
cdn1-C4 **(AF270425)**	TTCTATGATATGATTTTTTTAGG	Sense
	XXXXXXXXXXXXTTTTTAGGTGG	Sense
cdn1-C5 **(AY800106)**	TACAAATGCAATTGAGAGGTGG	Sense
	CCGACAAGAATATTGATGCTGAA	Anti-sense

PAM: Protospacer adjacent motif (highlighted in green); Intron sequence is highlighted in gray colour; X: Nucleotide sequence unknown but will be determined later from the whole genome sequence.

Thank you